

# Fair determination of student grades from marks

Author ACE32

Affiliation ACE32

Email@ACE32

## Abstract

Schools and departments in Australian universities assess students in their units through a number of assessment components. Often these are grouped as internal (or in-semester) components and an end of semester examination. Scores are added to derive a total that is then used to assign an assessed grade to the student. Typically, to ensure that students do not neglect either of the two main assessment components, policies define minimal performance standards for each component.

The paper examines common practices in making decisions at the Pass/Fail boundary, the mathematical fundamentals in the application of these policies and inherent in marks, and proposes that one of two processes be used involving error-estimation by the markers. It recommends that one of these approaches should be adopted in preference to the traditional one-mark process.

**Keywords:** Assessment, assessors meetings, grades, marks, scores, university grading, pass, fail, examination, assignment.

## 1 Introduction

Three years ago, end-of-semester School assessors meetings in the School of Computing, ACE32 University used to be marathon sessions. A unit lecturer-in-charge would introduce student marks and grade recommendations for their unit. Other assessors would closely scrutinise the recommendations against the School's published policy and its interpretation as applied to the other units during the meeting. Frequently, the meeting returned to an already discussed unit, to revise a grade in line with a fresh interpretation being discussed by the meeting.

The assessors came out of these long sessions unsure whether they had been consistent in their interpretations and students had been treated fairly in accordance with the published school policies. Since then the School has adopted uniform practices in the presentation of the recommendations. Amongst them is the practice of using a pre-programmed spreadsheet to compute scores before reviewing them against the published policy of the School. The meetings have become more manageable and their durations have become acceptable.

However, was the issue disturbing the assessors just the uniformity of the presentation styles? In this paper, it is argued that use of spreadsheets with pre-programmed macros is useful but has hidden some real issues that were at the root of the long meetings.

This paper has two goals. Firstly, it reviews some common and widely-used *ad hoc* practices regarding

discretization of the marks. It is shown that common practices fail to meet markers' intentions. It also examines issues related to the determination of grades from scores. Secondly, the paper is intended to initiate research into the manner in which present day computational tools can be used to assess students in a way more consistent with the limitations inherent in the marking process.

Section 2 introduces the published policies of our school, which are typical of most schools in Australia. Some examples are presented to describe the observed reasons for the difficulties noticed when the policies are put into practice by the assessors. The use of a spreadsheet implementing the policy through a uniform set of macros eliminates the observed difficulties. Section 3 takes a fresh view of the marking process and suggests a different approach for determining grades, reflecting the limitations of marking. In Section 4 some examples are used to compare the traditional approach against the proposed method. The issue of scaling of the marks and its fairness against the judgement made by the markers is also examined.

## 2 Background to the problem

To understand the contention that caused extended and never-ending rounds of arguments among the assessors, Appendix 1 reproduces the School of Computing, ACE32 University policy on assessments. Appendix 2 shows the policy in a pictorial form for easier comprehension. The published policy, however, does not lay any procedural arrangement through which the policy should be applied or the component marks combined. In other words it is outcome-directed rather than process-directed.

Each procedure to apply the policy to determine grades carries with it an intrinsic interpretation and algorithmic modification of the data. The different outcomes these interpretations produce is highlighted through a series of examples in the remainder of this section.

However, at this stage it would be appropriate to stress that the paper is about the computational processes underlying the determination of the grades from marks. The validity and nature of the assessments used to assign marks is outside the scope of this paper. Excellent literature exists to advice about the assessment practices appropriate to the maturity and age group of the students and purpose of the course (See, for example: AAIA website). In what follows, the reader will be assumed to be familiar with the basics of measurement theory. The measurement theory indicates the validity of transformations, operations and aggregations that can be applied to various measurements (Kan 1995, Pfleeger, 1998). We also make use of some basic statistical

formulae and definitions. Reader wishing to revise them can see any of a range of text from computer performance analysis classic (Jain, 1991) to a book on statistical methods (Hayslett, 1971). Finally, the paper also makes use of mathematical practices related to truncation and rounding of real numbers. An excellent coverage of the topic can be found in (Goldberg, 1991).

## 2.1 An example

Suppose a lecturer sets two assignments of each of equal weight ( $w_1, w_2, \sum w = 1$ ) as internal marks for the unit. Assignment 1 has a maximum  $mx_1$  of 75 marks and Assignment 2 a maximum  $mx_2$  of 125 marks. A student, Sarah, receives 30 marks ( $m_1$ ) in Assignment 1 and 49 ( $m_2$ ) in Assignment 2. Has Sarah met the 40% threshold ( $T$ ) necessary for a pass grade assessment?

In the rest of this sub-section different algorithms will be applied to answer the question. Assume that the internal marks contribute 30% to the final scores in the unit. The values  $w$  and  $T$  are rational fractions, and  $m$  and  $mx$  are integers. Real values are denoted by an overbar, eg  $\bar{m}$ .

### 2.1.1 Algorithm LCM

Marks in each assignment are translated to a new  $max$  which is the Lowest Common Multiple (LCM) of the two maximums, in this case 375, and then added to give a maximum of 750.

$$m = \frac{\left( \frac{m_1 \times mx_2}{GCD(mx_1, mx_2)} \right) + \left( \frac{m_2 \times mx_1}{GCD(mx_1, mx_2)} \right)}{2 \times LCM(mx_1, mx_2)}$$

Thus, Sarah's internal marks are reported as 297/750. Since  $40/100 = 300/750$  a student scoring 300 or more meets the stipulated 40% requirement, and Sarah fails to meet it. The scheme is easily modified to accommodate non-equal weights.

This algorithm is never used to present marks to assessors as any comparison across units becomes impractical due to different denominators appearing in different units. However, we can view it as the exact computation to which practical algorithms approximate.

### 2.1.2 Algorithm R

Algorithm R (Real) converts each internal mark to a real number and carries out the computation in real arithmetic. We assume that floating point arithmetic is sufficiently close to real arithmetic for the difference to be ignored.

$$\bar{m} = w_1 \frac{m_1}{mx_1} + w_2 \frac{m_2}{mx_2}$$

No rounding of the marks takes place until the final stage where the overall mark in the unit is set to an integer for transmission to a central university authority. In our example, the student receives an internal mark of 0.396. For convenience we often represent this as 39.6%. Not surprisingly Algorithm R also determines Sarah as having failed to meet the internal marks condition for a pass.

### 2.1.3 Algorithm RR

The algorithm (Real Rounded) is similar to Algorithm R except that the total marks are rounded to whole percentage points. Thus, 39.6% (0.396) would be written as 40% (0.40). In statistical practice if the percentage fraction is exactly 0.5%, to avoid bias the rounding algorithm selects the nearest even number; however assessors tend to accept 'up rounding' (the next larger integer) in this case as biased in favour of the student meeting the threshold.

### 2.1.4 Algorithm RW

Another algorithm (Real Weighted) translates scores to the weight in the final unit score. In real computation this score for Sarah is 11.88%. Repeating this for the threshold, a student requires a score of 12% for a pass. In real arithmetic the consequences are the same as algorithm R.

### 2.1.5 Algorithm RWR

This algorithm is same as Algorithm RW except the scores are rounded to the nearest percentage. However now the student passes, because their score (11.88%) is rounded to 12%, the threshold.

### 2.1.6 Strict and lax interpretations

All the algorithms except LCM have both a *strict* interpretation (as described) and a *lax* interpretation. In the latter a lower threshold  $T - \epsilon$  instead of  $T$  is used. For example Algorithm R takes  $T$  as 40% whereas a  $R_{lax}$  might take it as 39.5% ( $\epsilon = 0.5\%$ ). In other words, Algorithm  $R_{lax}$  is equivalent to R with  $T_{lax} = T - \epsilon$ .

Both strict and lax interpretations have their supporters; each with valid arguments. However a lax decision will vary depending on the agreed interpretation. The interpretations are unpublished and may vary on the day.

The reasons for adoption of lax algorithms are not mathematical, but are rooted in psychological and legal issues. Setting the actual threshold lower than the published one provides the lecturer with a ready defence when confronting an aggrieved student "But, your mark was *far* too low to pass." Anyone within coo-ee of a pass mark was passed. The decision is still due to an arbitrary cut-off, but practice and policy are divorced.

Fortunately for universities, few students pursue the legal angle: "Why did you pass XYZ, thus devaluing my degree (and the work I put in), when he/she didn't meet the published criteria?" This is probably due to the secrecy with which universities surround their raw assessment data, and unpublished procedures.

## 2.2 Algorithm comparison

The authors assert that rounded algorithms RR and RWR are hangovers from primitive manual computation to make arithmetic simple, and should never be used in present practice. They introduce unnecessary 'noise' (error) into the mark assembly process. It is not difficult to maintain accuracy to floating point precision right up

the point of decision at an assessor's meeting. This may be excessive for the accuracy of marks, but at least it does not introduce any unintended errors.

The authors also assert that Algorithm RW is undesirable (but not unworkable) as it makes the task of the assessor's meeting more difficult by changing the base of the component mark. The use of Algorithm R exclusively is recommended, and this is easy to program in a spreadsheet or a computer program.

### 2.3 Uniform standards

It is unfair that a student's result may vary due to the adoption of unpublished processes, sometimes idiosyncratic to the unit and lecturer. We speculate from the above examples that the apparent cause of lengthy discussions in the assessors meetings primarily stemmed from the varying formats in which the results were presented to the assessors, and procedural arguments.

In our School, a pre-programmed spreadsheet has been developed to provide a uniform presentation format for all unit assessment. Real numbers and real arithmetic are used uniformly (Algorithm R) though displayed results are rounded to 0.1 percentage point as adequate accuracy. Assignment marks and end of semester examinations are both presented as percentage scores, as are the final total scores. The spread-sheet also highlights students not meeting a criterion or threshold for a pass grade.

However, assessors during the meeting agree on an  $\epsilon$  below the published threshold to which they let the actual threshold drop. Thus, the actual practice used is an intermediate between the strict and lax interpretations. The uniformity has reduced the arguments. However, is the automated computation glossing over a real problem that lies underneath?

## 3 Analysis and a fresh look

### 3.1 Assumptions behind marking

The few basic axioms accepted by everyone for the marks in an assessment task of a unit are:

1. [Discretization] Marks are awarded on a discrete scale.
2. [Ordering] The possible marks are ordered.
3. [Directionality] A student receiving a mark later in the ordering has a higher achievement than a student receiving an earlier mark.
4. [Linearity] The steps between successive marks are equivalent.

To illustrate Assumption 4, the step between 39 and 40 is the same as between 84 and 85, and can be gained by an equivalent performance. This is easiest demonstrated in a multiple-choice test or a multi-question examination, but is nevertheless generally accepted over all assignments.

*Corollary:* Given axioms 1–4, each mark on an assessed component can be represented as a number and is usually denoted as a rational fraction of the selected mark over the maximum mark (often implicit). Arithmetic and relational operators are then defined over marks.

A further assertion is proposed that might be more contentious, but is logical and inescapable:

5. [Noise elimination] All rounding is undesirable in intermediate mark computations, as it introduces unnecessary quantization noise (in the engineering sense) into the process.

### 3.2 Algorithm RANGE

Depending on the assessed item, a student can only be awarded a discrete score. For example:

- In a multiple choice assessment where each question is worth 2 marks, only marks that are even numbers will be awarded.
- An item marked out of 10 where the marker can award  $\frac{1}{2}$  marks such as  $5\frac{1}{2}$  is nevertheless discrete and can be mapped on to integers out of 20.
- Some tasks which are marked out of 100 are arranged to award marks in multiples of 5.

Consider a student receiving mark  $m$  and let  $d$  = the step size of potential marks. Then the above assumptions determine that the student has achieved higher than students gaining  $m-d$  and lower than students gaining  $m+d$ . Further assume that the markers will concede that a marking error of  $d$  (one step) is possible.

In Algorithm RANGE, it is proposed to maintain records for each student of both  $m$  and these lower and upper bounds. In computation and record keeping, all three will be weighted, added etc. The net result is that each mark will be accompanied by an upper and lower bound due to the discrete nature of the marking scheme, abbreviated as UB and LB respectively.

To illustrate this suppose each of Sarah's two assignments were scored using criteria that assigned marks in units of 3 ( $d = 3$ ). The scores are calculated again using the three non-noisy algorithms (see Table 1).

**Table 1 – Algorithm RANGE**

Algorithm		L	R	RW
Assignment 1	LB	27	36.0%	5.40
	$m$	30	40.0%	6.00
	UB	33	44.0%	6.60
Assignment 2	LB	46	36.8%	5.52
	$m$	49	39.2%	5.88
	UB	52	41.6%	6.24
Total	LB	273	36.4%	10.92
	$m$	297	39.6%	11.88
	UB	321	42.8%	12.84
Threshold $T$		300	40.0%	12.00

Each algorithm indicates that Sarah has not been definitely found to be falling below the threshold of 40%

and therefore fairness requires that she be deemed ‘not failed’.

This technique eliminates the need for assessors to set an *ad hoc* band, since the band of potential marks is identified.

In practice the data storage might be simplified. It is not necessary to keep three marks unless non-linear scaling is employed. Retaining  $m$  and  $d$  is sufficient, and  $d$  is the same for all students and marks in an assignment task, so it only needs to be weighted and added in parallel with the mark set.

However, this is not exactly true. Students who fail to submit an assessment task have a zero mark with no error, up or down. Also students who achieve full marks (100%) in an assessment do not have the possibility of a higher mark. Taking account of these is a simple computational task but suggests that keeping all three values for each student is a reasonable approach.

```

LowerBound :=
  if (m-d < 0) → 0;
  else → m-d;
fi;
UpperBound :=
  if (m = 0) → 0;
  (m+d > mx) → mx;
  else → m+d;
fi;

```

The discrete steps of a subjectively marked assignment are a lower bound on the potential error, but it is also possible for markers to determine the possible error  $d$  in an assignment task, independent of the discrete scale. For example, a marker might nominate a probable error range of  $\pm 5$  for an assessment marked out of 100. Even in the case of a supposedly objective multiple-choice test, an assessor might consider that the questions might not be perfect, and admit to an error in the evaluation of student knowledge. Pursuing this track would extend the value of the technique considerably.

Note that RANGE is ‘best-case’ for the student. The student’s marks are given the *best* possible interpretation for *each* assessment object. Note also that no use is made in the example of the lower bound, due to the catastrophic consequences of a *Fail*. Were the process to be used to decide between awarding a *Credit* or *Distinction* grade, both bounds might be relevant.

### 3.3 Algorithm STATS

The second algorithm takes a statistical approach. Suppose that each assessment object has an error distribution (modified for extreme cases) associated with it. Such an error distribution is determined by the marker based on their certainty of the result. Except perhaps in the case of multiple choice assessments and maybe not even then, even the most confident marker will admit the possibility of a one-mark error, perhaps assigning probabilities to discrete marks as follows:

**Table 2 – Probability of Error**

Mark	Probability
$\leq m-2d$	0
$m-d$	0.25
$m$	0.5
$m+d$	0.25
$\geq m+2d$	0

Now the probabilities of the various marks can be combined as they are weighted and added. Assume that the potential errors in each assessment are uncorrelated. (The case of a common marker who consistently victimizes or favours a particular student must be treated in other ways.)

When combining scores, it would be possible to keep a variable-length list of all the possible scores with their associated probabilities. However, this would be difficult to incorporate into spreadsheet programming, and in the following example the number of possible scores would be  $3 \times 3 = 9$ . The number of possible scores can increase rapidly with the number of assessments.

A simpler method is to characterize each discrete probability distribution by a Gaussian error distribution, where the mark is the mean, and a variance is computed from the possible error data. For the error data in Table 2, the variance is  $\frac{1}{2} \times d^2$  and the standard deviation:

$$\bar{\sigma} = \frac{d}{\sqrt{2}} = 0.7071 \times d$$

For several distributions with means, variances and standard deviations  $\bar{m}_i, \bar{v}_i, \bar{\sigma}_i$ , the combined mean, variance and standard deviation are:

$$\begin{aligned}
\bar{m}_{combined} &= \sum \bar{m}_i \\
\bar{v}_{combined} &= \sum \bar{v}_i \\
\bar{\sigma}_{combined} &= \sqrt{\bar{v}_{combined}}
\end{aligned}$$

Each student’s mark then needs to be accompanied by only one additional piece of data to accompany each actual mark or combination mark. Although it is convenient to work with variances in the computations, the storage and display of standard deviations is better as they have the same scale as the accompanying mark. Note that the more components, the more the final distribution tends to be Gaussian anyway.

In Table 3, the scoring for Sarah has been recomputed. The probabilities in Table 2 are used, with a mark discrimination  $d=3$ . There is a new degree of freedom for a policy or assessors to choose: the level of acceptable uncertainty that the correct mark is above a computed bound. In the table uncertainty levels of 20% and 10% are used for illustrative purposes. Another alternative may be to present the probability that a correct mark passes the threshold.

**Table 3 – Algorithm STATS**

Raw data			Mean	Std Dev
Assignment 1	$d$	$\pm 3$	40.0%	2.83%
	$m$	30		
	$mx$	75		
Assignment 2	$d$	$\pm 3$	39.2%	1.70%
	$m$	49		
	$mx$	125		
Combined mark			39.6%	1.65%
0.2 (20%) uncertainty UB			41.0%	
0.1 (10%) uncertainty UB			41.7%	
Probability of mark $> T$			0.34	
Threshold $T$			40.0%	

Sarah passes, since she cannot be shown to be definitely below the threshold with a high degree of confidence. It would be desirable for a policy to set the appropriate level of probability/uncertainty prior to an assessors meeting.

### 3.4 Comparison and analysis

Algorithms RANGE and STATS are both computationally feasible in a spreadsheet. Both algorithms require the markers to provide an assessment of their possible marking error. Both algorithms can present data in a form readily understood by the assessors in an assessors meeting.

Algorithm STATS requires the assessors to have a rudimentary knowledge of basic statistics and probability, but the extent is not large and this is not considered an obstacle to its use.

The main difference is their treatment of the uncertainty, and their large scale behaviour:

- RANGE gives the student the best possible interpretation for every assignment, and thence the best possible mark, which may be improbable. STATS assumes that the errors in marking each assessment are uncorrelated, and thus produces a realistic view of the total error; a favourable interpretation can then be made at the final stage.
- The behaviour differs with many components. Consider a unit with  $n$  marked components, each with the same  $mx$  and  $d$  or  $\sigma$ . Combining these with RANGE and STATS yield:

$$d_{\text{RANGE}} = \frac{n \times d}{n} = d$$

$$\bar{\sigma}_{\text{STATS}} = \frac{\sqrt{n \times \bar{\sigma}^2}}{n} = \frac{\bar{\sigma}}{\sqrt{n}}$$

In other words with more components the standard deviation of STATS will reduce slowly to reflect less likelihood of error, whereas RANGE maintains its error range. This is particularly relevant if the algorithms were to be

applied to marking of individual questions in an examination paper as assessment objects.

Note that both algorithms are capable of having the applicable treatment rules coded into a spreadsheet to make the adjustments automatically (or to recommend them to the assessors meeting).

The focus in this paper has been on using the algorithms to deal with the Pass/Fail decision region. The same algorithms can be applied to decisions at other boundaries such as Credit/Distinction, but the decision is not so one-sidedly biased. A subsequent paper on making these decisions is foreshadowed.

## 4 Test cases and scaling

### 4.1 Effect of discrimination factor $d$

It is possible for the grade from our range process to show an anomaly from conventional wisdom across different units. Consider two students John and Marie in two separate units with equally weighted internal assessments. Marie has apparently scored slightly higher than John, though in a different unit. Table 4 shows the raw data analysed using RANGE. Similar results would be produced by STATS.

**Table 4 – Effect of discrimination**

Name	Assessed Item	Marks	Maximum score	$d$
John	Assign 1	38	100	2
	Assign 2	20	50	5
	Assign 3	75	200	5
	<b>Total</b>	<b>38.5</b>	<b>100</b>	
Marie	Assign 1	19	50	0.5
	Assign 2	40	100	0.5
	<b>Total</b>	<b>39</b>	<b>100</b>	

If the total marks in the table above comprise all the internal marks, and neither student meets the apparent 40% requirement stipulated in the School policy, there is still a difference in how markers have marked them. Using the range process for simplicity, John has not been determined to be definitely below the threshold since  $((38+2)*2+(20+5)*4+(75+5))/6 > 40$ . On the other hand, Marie has been determined to be definitely below the threshold:  $((19+0.5)*2+(40+0.5))/2 < 40$ . John passes, Marie fails. The statistical approach would be similar.

The discrimination of a mark in an assessed component is affected by how the component is broken into smaller units for the purpose of assessment, multiple markers, the marking scheme, and on many other factors. In Marie's unit the relative discrimination  $d$  was set much lower than for John's unit, bordering on infeasible accuracy.

A component that is open to cheating is also probably less precise in its assessment of the students. Markers should be asked to provide judgment in relation to the correct-

ness of their assessed marks. With experience, markers should become more expert in assessing discrimination.

## 4.2 Using bounds to limit scaling

Linearly scaling of marks for one or more assessed components should not be allowed as it violates the legal requirement of keeping the documented weights for the assessment components (scaling being mathematically equivalent to a re-weighting). However, for various reasons, assessors sometimes resort to scaling the marks.

Scaled scores are always a contentious issue, but can be limited within the framework presented in this paper. Firstly, it is only appropriate to scale the overall total marks as this does not discriminate against any particular assessment component. Secondly, scaled marks should not be allowed to push too many students outside a band consistent with the markers' discrimination determinations.

An appropriate extent to which scaling of the marks is allowable can be shown by an example. Three students in the unit are Ann, Bert and Cyd. The three assessment items in the unit are weighted 25%, 35% and 40% respectively. Suppose the markers believe that the first assessment has an error band of 3% either way; error in the second assessment may range from -5 to +3 marks, and the last assessment has multiple markers who may have used different standards leading to an estimated error of about 7%.

**Table 5 – Effects of Scaling**

Item	Max marks	Ann	Bert	Cyd
Assign 1	20	10	15	5
Assign 2	50	30	23	10
Assign 3	100	60	60	70
LB		51.95	53.11	35.60
<b>Total</b>	100	<b>57.50</b>	<b>58.85</b>	<b>41.25</b>
UB		61.66	63.19	45.50

Should the marks be scaled upward by 7.2%, Ann's score will be just below 61.66, the maximum predetermined by the marker. Likewise, marks could be scaled down by 9.65%. Scaling by bigger factors would be inconsistent with the bands set by the markers and would penalise students.

Again, similar limits to scaling could be set by STATS, but the limits are fuzzy and determined by acceptable confidence levels. A scaling that moved a mark by  $\sigma$  might be acceptable, but  $3\sigma$  would not without declaring that the assessment process was faulty.

## 5 Conclusions

The paper has reviewed methods of transmitting markers' assessments to a board of assessors. It challenges the

common practices of rounding marks (for example, 39.51 is changed to 40 while 39.49 becomes 39) as being invalid and inconsistent with the marking processes. It has also established that today's computing technology is readily available to implement better processes, with marginal extra work for markers. Each assessment object mark is associated with a band of scores or a distribution of scores rather than a single number. The range/distribution of scores depends on factors that the markers determine for their assessment. This data is then carried through in spreadsheet computations to arrive at a consolidated band or distribution, a decision, and an eventual result.

The next step in this research is to test the RANGE and STATS algorithms against historical data, say for the last semester. The markers are still available and can be asked to estimate their errors retrospectively. Unfortunately, trying to test the process prospectively would be likely to perturb the assessment process simply through the effect of the intervention.

It is time that the computing community re-examines the methods that are used to compute an overall assessment for its students. If the computing community doesn't do it, who will show the way?

## 6 References

- AAIA: *Primary assessment practice: evaluation and development material*. Association of Assessment Inspectors and Advisors. <http://www.aaia.org.uk/>. Accessed 24 Aug 2004.
- Goldberg, D. (1991): *What Every Computer Scientist Should Know About Floating-Point Arithmetic*. ACM Computing Surveys Vol. 23(1): 5-48.
- Hayslett, H.T. (1971): *Statistics: made simple*. London, W.H. Allen & Co., Ltd.
- Jain, R. (1991): *The art of computer systems performance analysis: techniques for experimental design, measurement, simulation, and modeling*. New York: John Wiley & Sons, Inc.
- Kan, S.H. (1995): *Metrics and models in software quality engineering*. Reading, Ma.: Addison-Wesley Publ. Co.
- Pfleeger, S.L. (1998): *Software engineering: theory and practice*. NJ: Prentice-Hall International, Inc.

## Appendix 1: Copy of School Assessment Policy

### Guidelines for Assessment

The aim of these guidelines is to ensure the fair and equitable assessment of students enrolled in all units offered by the School of Computing. These Guidelines are based on the Assessment Guidelines approved by the Faculty of Science and Engineering.

These guidelines will be used for all units offered by the School of Computing, unless stated otherwise in unit outlines and approved by the Head of School.

### Attendance Requirements

Students are required to attend a minimum of two-thirds of lectures, tutorials and laboratory classes. A higher attendance rate can be imposed but this must be stated as part of the written notification of assessment provided to all students.

A student who does not meet the minimum attendance requirements may be declared not eligible (NE) to be assessed for the unit by the Head of School. However, declaring a student not eligible on the grounds of attendance should only be done if the lecturer/tutor has maintained attendance records.

### Examination Grades

#### *Grades of pass*

The School uses the AV-CC's guidelines for grades of pass:

Pass (PP):	50–59%
Credit (CR):	60–69%
Distinction (DN):	70–79%
High Distinction (HD):	80%+

Overall assessment will be based on the student's performance throughout the semester (in semester) as well as in a formal examination.

In order to achieve a pass (or better) result, a student must obtain at least 40% of the mark in each assessment component (in semester & formal examination) and 50% in the combined total.

Assessors and the School Moderation Committee should pay close attention to borderline results to ensure that the student is awarded the appropriate grade. It is legitimate to report such borderline marks provided that they have been reviewed.

#### *Supplementary and terminating grades*

Students gaining 45% - 49%, with at least 35% of the mark in each assessment component, would be granted a terminating pass (TP) grade.

Students gaining 45% - 49%, with at least 40% of the mark in each assessment component, would be

considered for recommendation for a terminating pass with permission to enter a supplementary (TS) grade.

Students gaining 50% or better, with at least 35% of the mark in each assessment component, would be considered for recommendation for a terminating pass with permission to enter a supplementary (TS) grade. Such a TS grade will be used without a mark.

Recommendation for a TS grade should only be made if the assessors are confident that, with extra work, the student could achieve the level required of a full pass in the unit.

Where a student obtains 45% or better in the combined total but achieves less than the minimum 35% in either assessment component, NN (or NS) grade will be used without a mark being recorded in order to differentiate this group of students.

The maximum mark awarded to a student who fails the unit (NN) will be 44.

The fail with permission to enter a supplementary (NS) grade should only be used for special circumstances.

The awarding of supplementary grades (TS, NS) is done by the Faculty's Assessment Committee. Schools make recommendations for these grades which may be accepted or rejected. Generally, a student will have to have gained full passes in at least 50% of the units attempted to be given access to the supplementary examination system.

If a student fails the final unit to complete a degree, a supplementary grade (TS or NS) would be recommended.

#### *Withheld*

The withheld (WT) grade may be used if there are genuine reasons why a student has been allowed an extended deadline for assignment work. All WT grades must be finalised and the actual grade reported by the beginning of the supplementary examination period.

### Deferred Ordinary (DO) Examinations

A deferred ordinary (DO) examination may only be granted by the University Examinations Officer acting under the authority of the Academic Registrar. There are approved procedures and precedents for the granting of DOs.

A student granted a deferred ordinary examination in a unit has no access to the supplementary examination system for that unit.

### PDF Version of this Document

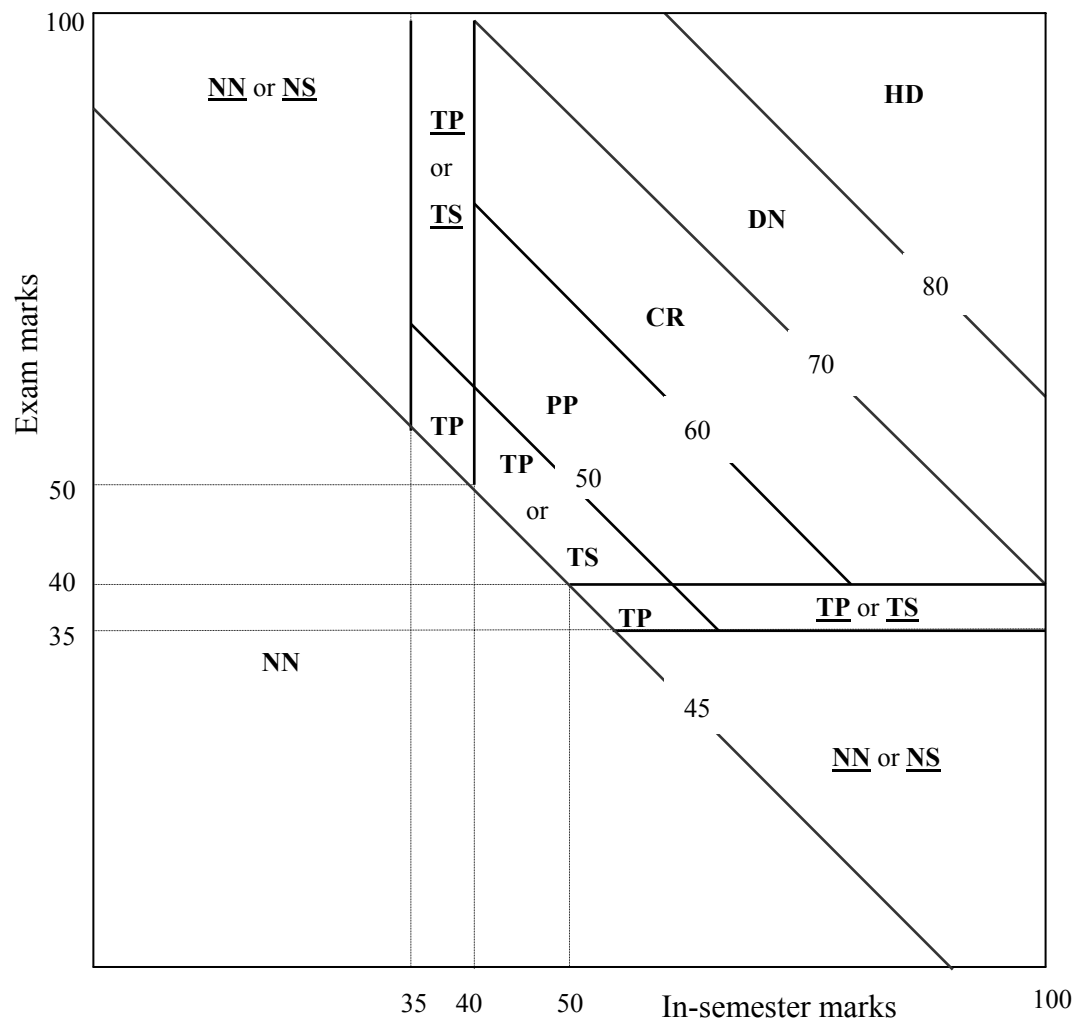
A pdf version of this document is available at <http://URL.ACE32>.

(Signature)  
Head, School of Computing

February 2001



## Appendix 2 – The policy graphically



**Figure 1:** Regions for various grades defined by in-semester and end-of semester examination marks. The figure is drawn based on in-semester component weighted as 50% and end-of semester examination as 50%. An underlined grade indicates that a mark is not recorded as it would be misleading.